Draft

Proposed Procedures for Estimating the Limit of Detection

Consensus Group
Committee I on Detection

for

Proposal to USEPA for Replacement of
40 CFR, Part 136
Appendix B MDL Procedure

Procedure Number:
Revision Number: 0

Date Initiated: 9/8/2004   Date Revised: 6/30/2005
Effective Date: DRAFT 9   Pages Revised: See Highlights

Total Number of Pages: 5

_____

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 2 of 24	Date Initiated:

0.0 SCOPE

To provide a procedure by which an individual laboratory may derive accurate estimates of routine method sensitivity for most analytical procedures for which they report results. Short term, long term and on-going estimates for L_C (Critical Value) and L_D (Detection Limit) are derived for uncensored methods as well as censored 2D and 3D methods. The sensitivity estimates are then used as benchmarks for accurate reporting of analytical test results which fall below the LQ (Limit of Quantitation).

1.0 INTRODUCTION

1.1 The Method Detection Limit (MDL) procedure as specified in 40 CFR Part 136 Appendix B defines the detection level as that level where one can be 99% certain that if the test method results indicate the presence of the analyte being evaluated (1% false positive error rate). Unfortunately, the MDL does not satisfy this definition in actual practice, because it is a poor estimate which does not take into account long term variability (*e.g.,* varying instrument conditions), blank contamination or recovery bias. The MDL in theory is equivalent to the Critical Level (L_C) not the Limit of Detection (L_D). The following procedure rectifyies these shortcomings of the MDL and provides a practical means to produce accurate and realistic estimates for L_C and L_D when needed,

1.2 The reporting of data below the Limit of Quantification (L_Q) is not always necessary, therefore estimates for L_C and L_D are only required when data are reported below the L_Q. The need for reporting down to L_C and L_D for specific targets should be defined in the Data Quality Objectives for the project before it is intiated. The short term estimates for L_C and L_D as described in this procedure could also be used to demonstrate proficiency for the initial start up of a method. Long term estimates for L_C and L_D as described in this procedure will automatically replace the short term estimates.

1.3 This procedure is potentially applicable to analytical measurements that produce at least ratio scale data. Detection limits are not meaningful for nominal, ordinal or interval measurement scale data. A ratio level measurement scale has all the properties of an interval level scale plus a meaningful zero point. For example, temperature readings in degrees F and degrees C are interval level data ($20^{o}C$ is not twice as "hot" as $10^{o}C$), therefore detection limits would not be determined for F or C thermometers. It should also be noted that although this procedure applies to test methods capable of producing ratio scale data, it may not be necessary to determine the critical value and detection limit of a test method; the need to do so is inherently a function of the end use of the data. For example, if the quantitative range of a test method has been established, it would not typically be necessary to determine the critical value and detection limit if the test method were being exclusively used to monitor changes in the measurement variable that all fall well within the quantitative range of the method.

1.4 The Detection Limit L_D, only assures the presence of the analyte with confidence, however the level of uncertainty is unknown at the L_D for a single measurement. Numberical values are not quantitative at the L_D. The numberical value of a single measurement is only known with a known level of certainty at the L_Q.

1.5 The distribution of data in the region of detection is generally assumed to be normal (Gaussian), therefore that assumption is also made for this procedure. If the user suspects that their data is non-normally distributed (Log-Normal for example), we recommend several techniques to evaluate its distribution for normality. If data is log-normally distributed, it is possible to perform a log transformation of the data and then apply the normal distribution techniques to the transformed data. A discussion of dealing with data which is neither normally nor log-normally distributed is beyond the scope of this procedure.

We recommend using the following tools to evaluate data distribution when you have at least 20 data points. Summary statistics such as the coefficient of variation or coefficient of skewness may be used, to give an indication that your data may not be normally distributed. Graphical techniques such as probability

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 3 of 24	Date Initiated:

plots and box plots may be used to evaluate whether your data may be normally or log-normally distributed. All of the aforementioned techniques must be used with censored data. Formal tests to evaluate the normality of your data include the Shapiro-Wilk test (n≤50), the Sapiro-Fancia test (n>50) or D'Agostio's test (n>50).

Shapiro, S.S. and M.B. Wilk, 1965. An analysis of variance test for normality (complete samples). *Biometrika,* 52:591-611.
Shapiro, S.S. and R.S. Francia, 1972. An approximate analysis of variance test for normality. *Journal of the American Statistical Association,* 63:1343-1372.
D'Agostino, R.B. 1971. An omnibus test of normality for moderate and large size samples. *Biometrika,* 58:341-348.

2.0   SUMMARY OF PROCEDURE

2.1   The specific procedure followed depends on the type of analytical technique being employed. For methods that typically produce a numerical result for a method blank (Uncensored Method), the results of between 7 and 20 blanks are used to statistically calculate a Critical Value and Limit of Detection. For methods that typically do not produce a method blank with a numerical value (Censored Method), an estimate of the Limit of Detection is derived based on the instrument "noise" level present in a method blank. The estimate is then tested using spiked blanks, and if necessary, adjusted. Seven spiked replicates are then analyzed and a Critical Value calculated. The qualitative capabilities of the analytical method determine which Censored Method procedure will be followed. For a 2D method, the Limit of Detection is set based on the quantitated value of the seven replicates. For a 3D method, the Limit of Detection is determined by both the quantitative and the qualitative results of the seven replicates.

2.2   The initial Critical Value and Limit of Detection results are initially based on very small data sets. Once enough data have been generated, the short term estimates of these values must be replaced by long term estimates. If the laboratory already has enough data to generate long term estimates the derivation of initial short term estimate may be skipped entirely. It should be noted that the initial, long term and on-going estimates for L_C and L_D are all valid estimates. However, as estimate which are more representative of the laboratory process are generated they will replace the previous estimate. For example, once a long term estimate is generated it will replace but not invalidated the initial estimate and on-going estimates will update but not invalidate long term estimates.

2.3   The Critical Value and Limit of Detection are verified periodically (at least annually) by the monitoring of method blanks and the analysis of false negative quality control check samples (FNQSs) for both censored and uncensored methods. If the Critical Value and Limit of Detection can not be verified, their estimates are revised appropriately to provide a continuous improvement in detection estimates.

3.0   DEFINITIONS

Uncensored Method
Analytical methods that nearly always (at least 85% of the time) produce numerical values for method blanks (e.g., spectroscopic tests such as ICP-OES) are referred to as Uncensored Methods.

Censored Method
Analytical methods that frequently (> 15% of the time) produce non-numerical results for blanks (e.g., chromatographic methods such as GC and LC methods) are referred to as Censored Methods. Censored Methods are additionally categorized as "two dimensional" (2D) and "three dimensional" (3D) techniques. Censored methods sometimes yield a near constant blank signal or a flat numerical result.

2D Method

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 4 of 24	Date Initiated:

A 2D technique qualitatively identifies an analyte based on both a time dimension (retention time) and a response dimension (FID, ECD, TIC, UV, IR …). Dual column confirmation, wavelength ratios, or dual detector methods, while providing additional conformational information are still considered 2D techniques because they still measure only in the time and response dimensions.

3D Method
Relative to a 2D qualitative technique, a 3D technique provides additional qualitative information to identify the analyte of interest. The additional dimension of information includes, mass spectral information, FTIR spectral data and pattern recognition for the identification of multi-component analytes (e.g. PCB Aroclors by GC/ECD).

Critical Value, L_C and L_c
The Critical Level, L_C, is defined as the smallest amount or concentration of analyte that can be distinguished from a blank or zero at a high level of confidence; it is the smallest value at which a detection can be observed reliably. Conceptually, L_C is the smallest concentration that protects against false positives (Type I error) at a 1% error rate. This means that for every 100 measurements of a true blank only one value should fall above the L_C. While L_C represents the true Critical Level, L_c represents an estimate of the true Critial Level. For methods capable of reporting uncensored results, users may request that results less than L_c be reported. Detections may be reported to the critical value L_c but must be flagged as estimates when less than the Lq. Results less than the L_c will be reported as non-detect or as "< L_d" (e.g., or "L_d U," where the "U" qualifier indicates a non-detect).

Detection Limit, L_D and L_d
The Limit of Detection, also referred to as the detection limit. It is the smallest concentration that protects against false negatives (i.e. the smallest amount or concentration of analyte that must be present in a sample in order to be detected at a high level of confidence). At the L_D the Type II (false negative) error rate is 1%. While L_D represents the true Limit of Detection, L_d represents an estimate of the true Limit of Detection. L_d constitutes the lowest possible limit for the confident reporting of a non-detect. L_d is the smallest concentration that confidently results in measured values greater than L_c (or, for 3D methods, satisfies all method-specified identification criteria).

ISO
International Standards Orgainization

IUPAC
International Union of Pure and Applied Chemist

Limit of Quantitation, L_Q and L_q
Limit of Quantitation, L_Q, is defined as the lowest concentration that, in the context of some level of precision and bias, meets all method identification criteria and produces quantitatively reliable results for the end use of the data. While L_Q represents the true Limit of Quantiation, L_q represents an estimate of the true Limit of Quantitation. If detections less than the quantitation limit L_q are required to be reported, then detections greater than or equal to the L_c but less than L_Q must be reported as quantitatively estimated values.

MDL
Method Detection Limit as previously defined by the USEPA in 40 CFR, Part 136 Appendix B. The MDL is a single laboratory short term estimate of L_C, where $\alpha = 0.01$ following the "t" distribution and is typically based on six degrees of freedom.

Method Blank
An unspiked or non-fortified reagent water sample which proceeds through the entire method, including all preparatory and determinative steps.

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 5 of 24	Date Initiated:

FNQS

The false negative quality control sample (FNQS) is a method blank (e.g., reagent water) or "clean" sample that is spiked at (or near) L_D with the analyte of interest and processed through the entire analytical procedure to verify that such a spike will produce a detection.

4.0   PROCEDURE

4.1   The procedure followed depends on whether or not the analytical test generates method blanks with numerical values. Methods that typically generate numerical values for method blanks, such as spectroscopic test, will use the Uncensored Method (section 4.1.1). Methods that typically do not generate numerical values for method blanks, such as gas and liquid chromatography procedures, will use the Censored Method (section 4.1.2). Alternatively, the Censored Methods procedure (4.1.2) may be used even when the analytical test generates numerical results for method blanks, but this may result in larger estimates of L_C and L_D. Refer to Attachment 8.1 which depicts the procedure in a schematic flow diagram.

    4.1.1   <u>Uncensored Methods</u>

        4.1.1.1   This procedure is used if at least 85% of the method blank analytes are reported as numerical values. For initial demonstration of performance (e.g., for a new analytical method), collect results for method blanks generated during routine operation of the method. The method blank must go through all the preparation and analysis steps of the method. A minimum of 7 method blank results is required in order to calculate an initial estimate of the detection limit. Each method blank should be processed in a different preparation batch.

            NOTE: If it is necessary to initiate analysis immediately, an initial estimate of the critical value and detection limit may be made by analyzing seven blanks in a single batch. However, this short-term determination may underestimate routine variability. Replace the short term estimates by those determined using method blanks in a minimum of seven different batches as soon as they are available.

            A method is defined as a unique combination of preperative and determinative steps. For example, if estimates of L_C and L_D are derived for method 6010 (ICP/AES) using prep method 3010 (acid reflux digestion), separate estimates of L_C and L_D would be required if prep method 3015 (microwave assisted digestion) is substituted for prep method 3010.

        4.1.1.2   If a larger number of method blanks are available, then they should be used to estimate L_C and L_D as long as there is no reason to suspect that a change in method sensitivity occurred during the time period in which the method blank data were collected. The objective is to accurately estimate the "true" (population) standard deviation, σ. The larger the number of blanks used to calculate the sample standard deviation, *s*, the better *s* estimates σ (i.e., the smaller the factor K). However, it is not necessary to use more than several hundred (most recent) data points, since additional improvements in the accuracy of the estimate σ will be small relative to the data gathering and computational effort. If the laboratory has existing method blanks results available, for a current method, they may be used for the detection limit determination, bypassing section 4.1.1 and proceed directly to section 4.2 of this procedure.

            NOTE: This procedure will only be appropriate if numerical values are reported for each replicate blank used in the estimate. No more than

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 6 of 24	Date Initiated:

15% percent of the data set may contain blanks with non-numerical results and a minimum of seven numerical replicate measurements must be used in the calculation. Note that it is acceptable (and expected) that some method blank results will be negative values.

4.1.1.3   Method blank values <u>known</u> to be spurious errors that occurred during analysis should be discarded at the time of analysis, or where appropriate, corrected. The data set should consist of method blanks for which contamination was sufficiently small for the reporting of valid results to the data users. It is acceptable to apply statistical outlier tests, for example the Grubbs test (Attachment 8.3) to identify and discard anomalous method blank values for large historic data sets. However, for a small data set (e.g., n < 15) or for method blank data that have been recently acquired, a result should not be rejected solely on the basis of a statistical outlier test – a physical rationale for rejecting the result must be documented. Documentation of a physical rationale may be impractical for a large historical data set (e.g., n > 100) acquired over a long period of time. Under these circumstances, data may be rejected solely on the basis of statistical outlier tests, but this should be done with caution. The excessive rejection of method data will result in calculated detection limits and critical values that are biased low. Under no circumstances should more that 15% of method blanks be rejected.

4.1.1.4   Calculate the sample standard deviation, *s*, of the set of n method blank measurements. Calculate the critical value and the detection limit using the following equations:[1, 2, 3]

$$L_c = |\overline{X}| + s\,K_{\gamma,\,p,\,\nu}$$

$$L_d = |\overline{X}| + 2\,s\,K_{\gamma,\,p,\,\nu}$$

4.1.1.5   When calculating detection limit and critical value estimates several digits should be retained to prevent rounding errors. The final result should be rounded to a single digit using conventional rounding rules. The Limit of Detection is the highest concentration at which there are zero significant digits, therefore only the first uncertain digit is reported. Table 1 lists values for $K_{\gamma,\,p,\,\nu}$ for p = 0.99 and γ = 0.01 for various values of the degrees of freedom ν. For a set of n replicates, ν = n -1. Ideally (especially for analysis of inorganic analytes by spectroscopic methods), the mean of the set of n blank replicates should be near zero or much smaller than $s\,K_{\gamma,\,p,\,\nu}$.[4] A large positive sample mean that is significantly different from zero may be indicative of excessive blank contamination, inaccurate interelement correction factors or improper initial calibration. Method blanks that were rejected due to high levels of contamination should not be included in the data set. Otherwise, include ALL blanks that are representative of routine laboratory operations in the data set used to calculate L_c. This means a blank should not be excluded from the dataset unless the entire batch of data is also rejected. Method blanks should be run amoung and as normal samples. For example; if you don't normally proceed your samples by a blank or wash you should not do so with the method blank.

4.1.1.6   If the sample mean is significantly less than zero, then investigate the method (e.g., instrument calibration) to determine if the negative bias can be corrected. The magnitude of the negative bias should be minimized to the extent that is

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 7 of 24	Date Initiated:

practical ($|\overline{X}| < s\,K_{\gamma,\,p,\,\nu}$). However, if the mean is found to be significantly different from zero it must be used in the equations to calculate L_c and L_d.

4.1.1.7   If multiple instruments are to be used for the same test and it is desirable to report a single detection limit and critical value for the set of instruments, then a minimum of 7 method blanks must be analyzed for each instrument and the standard deviation calculated for each instrument. For a set of r instruments, the standard deviation for the ith instrument, s_i (where i > 1, 2, …r), is calculated using $n_i \geq 7$ replicates analyzed on that instrument. If all the instruments possess similar sensitivities, then a pooled standard deviation, s_p, may be calculated as follows:

$$s_p = \sqrt{\frac{(n_1 - 1)\,s_1^2 + (n_2 - 1)\,s_2^2 + \ldots + s_r^2\,(n_r - 1)}{(n_1 + n_2 + \ldots + n_r - r)}}$$

4.1.1.8   When the standard deviation for each instrument is determined using a set of only seven replicates, the standard deviations may be pooled when the largest standard deviation is no greater than two (for single analyte methods) to three (for multi analyte methods) times the smallest standard deviation. Alternatively, a F-test may be performed using the largest and smallest standard deviations to determine whether the standard deviations may be pooled.[5] Note that if the same number of replicates, n, is used for each instrument, the pooled standard deviation is simply the square root of the mean variance for the r instruments:

$$s_p = \sqrt{\frac{s_1^2 + s_2^2 + \ldots + s_r^2}{r}}$$

The pooled standard deviation (s_p) is substituted for *s* and the "degrees of freedom" $\nu = n_1 + n_2 + n_r - r$ for the determination of L_c and L_d.

4.1.1.9   If multiple instruments are used but the instruments possess significantly different sensitivities (i.e., the standard deviations are significantly different from one another), then establish separate detection critical values and detection limits for each instrument . Alternatively subsets of instrument with similar sensitivities may be pooled. As opposed to calculating a pooled L_c and L_d for a set of similar instruments, the highest critical value and detection limit may be used for the set of instruments , if this meets the measurement quality objectives for sensitivity.

4.1.2   Censored Methods

4.1.2.1   This procedure applies to analytical methods that often do not produce numerical values for blanks; that is, for methods that would produce non-numerical results for more than 15% of the method blank analyses. Use spiked replicates to generate sufficient data for censored methods. If the lowest possible detection limits are required, the level of the spike is critical– it needs to be high enough for reliable qualitative identification, but no higher, since variance usually increases with increasing concentration. . Censored methods are additionally categorized as "two dimensional" (2D) and "three dimensional" (3D) techniques as defined in Section 3.

NOTE: A method is defined as a unique combination of preperative and determinative steps. For example, if estimates of L_C and L_D are derived for method 8270 (semi-volatile GC/MS) using prep method 3510 (seperatory funnel liquid/liquid extraction), separate estimates of L_C and L_D would be required if prep method 3520 (conintuous liquid/liquid extraction) is substituted for prep method 3510.

4.1.2.2 Estimate the smallest analyte concentration that will produce a detected result and fortify a method blank at this concentration. As the term is used here, "detected" result is defined as a numerical value from a measurable analyte signal that is clearly distinguishable from background "noise" (i.e., the signal from a blank) under the routine operating conditions of the method which meets all qualitative method-specified identification criteria [6]. Typically, for methods that are readily amenable to evaluations of this nature, this concentration should produce an apparent analyte signal that is three to five times the apparent noise level (*e.g.*, via a qualitative visual examination of chromatograms for GC and LC methods). Specific qualitative or quantitative criteria for analyte identification are usually required for 3D methods (ranges within ion abundance ratios must fall for mass spectroscopy methods) and may be required, though to a less extent, for 2D methods (e.g., the detection of the analyte peak within a specified rention time windows for the primary and confirmatory columns). When specified by a test method, the measured result produced at the selected spiking concentration must be sufficiently high to meet all the method-required criteria required for analyte identification. The laboratory may use prior experience (e.g., prior analytical data) or consideration of the signal to noise to determine this estimate.

4.1.2.3 The estimate must now be tested and verified. Analyze at least a single spiked blank at the estimated lowest concentration of reliable qualitative identification through the entire analytical procedure (including all preparatory and determinative steps). If the analyte is not "detected" repeat the test at twice the original concentration used in section 4.1.2.2.

4.1.2.3.1 If the analyte is "detected" and the laboratory needs to demonstrate the ability to detect at a smaller concentration then the test may be repeated at half the original concentration used in section 4.1 2.2.

4.1.2.3.2 Note the smallest concentration, *x*, at which "detection" was achieved.[7] The laboratory may analyze several different spike levels at the same time in order to quickly determine the smallest concentration providing "detection." The spiking concentration for each instrumental system should be determined in this manner.

4.1.2.3.3 If multiple instruments are to be used to perform the same test and the user desires to use the same detection limit for all instruments, then the test of the detection limit estimate must be performed on each instrument, and the largest value of *x* from all the instruments may be used as the estimate.

4.1.2.4 Process (the preparatory portion of the method) a minimum of 7 replicates, each spiked at *x* (section 4.2.3.2). It is preferable to process the spikes in different analytical batches (e.g., to reduce the probability of under-estimating method variability). Analyze (the determinative portion of the method) the replicates on

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 9 of 24	Date Initiated:

each instrument. Existing data, such as the variability estimate (s) from low-level spikes previously analyzed, may be included.

4.1.2.5   Separate critical values and detection limits should be calculated for each instrument. However, if the sensitivities of the instruments are similar (i.e., for the determinative portion of the method), then a pooled standard deviation may be calculated as discussed in sections 4.1.1.7 through 4.1.1.9 and a single critical value and detection limit may be established for the set of instruments.

4.1.2.6   Estimate the critical level from the equation:

$$L_c = s\,K_{\gamma,\,p,\,\nu}$$

Verify that all the replicate spikes at the selected concentration (*x*) produce measured concentrations greater than the calculated critical value and any method-specified ID criteria are satisfied.

4.1.2.6.1   For censored 2D methods, if one or more of the $n \geq 7$ replicates (spiked at concentration *x*) fail to produce a measured value greater than L_c, the spike concentration should be increased until measured values greater than L_c can be consistently obtained. (A measured value greater than L_c is defined as a detection for 2D techniques.)

4.1.2.6.2   For censored 3D methods, if one or more of the $n \geq 7$ replicates fails to produce a result that satisfies all method-specified identification criteria, increase the spiking concentration until this occurs. It is not necessary to increase the spiking concentration to obtain a measured value greater than L_c unless the qualitative information provided by the method appears to be inadequate to confidently report the presence of the analyte. For example, a response greater than L_c would be required for a MS method when a target analyte did not possess any secondary ions or if only the Total Ion Chromatogram (TIC) response was being used, because in these examples the MS is being used as a 2D technique.

4.1.2.6.3   For both 2D and 3D analytical method, if a detection is not obtained for one or more replicates, analyze at least two replicates at the higher spike concentration and verify these replicates produce detections (e.g., measured results are greater than the calculated critical value L_c). Establish this concentration as the detection limit, L_d. The estimated detection limit, L_d, is the lowest spiking concentration that consistently produces a detection (i.e., results greater than the calculated critical value or fulfillment of all method-specified analyte identification criteria).

4.1.2.7   Blank Check – There may be some analytes in certain methods that are frequently detected in method blanks (e.g., common laboratory contaminants, such as dichloromethane or acetone for VOCs) despite efforts to eliminate blank contamination. As appropriate for these analytes, establish a critical value and detection limit based on the method blanks as described in section 4.1.1. If numeric values are not available at least 85% of the time or an inadequate number of blank results are available, use the procedure in section 4.1.2. Optionally,

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 10 of 24	Date Initiated:

estimates may be calculated using both procedures (section 4.1.1 and section 4.1.2) and the greater of the two values utilized.

4.2   Development of Long-Term Estimates of L_C and L_D

    4.2.1   The short-term initial detection limit estimations performed according to sections 4.1.1 and 4.1.2 use small data sets (n<20), and they <u>must</u> be replaced with estimates using larger data sets once the data is available. Sections 4.1.1 and 4.1.2 may be skipped if larger data sets (n≥20) are initially available.

    4.2.2   The initial estimate is replaced once 20-100 data points are available. If data is being pooled from multiple instruments of similar sensitivity following the sampe analytical procedure the combined number of data points available must be in the 20-100 range. If at least 20 points are not available one year after the initial demonstration of sensitivity was performed, repeat the calculation of L_c using the available additional method blank or FNQS data collected during this time. A total of at least 20 (ideally 100) points must ultimately be available to determine long-term estimates of L_c and L_d according to the equations and procedures specified in sections 4.1.1 and 4.1.2. *Data should not be pooled if it is suspected that method sensitivity has significantly changed during the period of time over which data collection was performed.* In particular, do not pool the data if the FNQS analyzed as part of the on-going demonstration of method performance failed to give rise to detections or the F test discussed below indicates method sensitivity has changed (section 6.2). Verify the new value of L_c using at least one spike at L_d as discussed in section 4.1.2.4.

    4.2.3   The data set used must be representative of routine analysis. For example, if the detection limits are to be determined using method blank data then all method blanks must be included unless the laboratory noted an unusual contamination or analysis problem and rejected the samples in the associated batch. Limited outlier removal using standard procedures such as the Grubbs test may be performed. A method blank result greater than the initial estimate of L_C is not sufficient evidence per se to remove the blank result from the data set used for the long-tern L_C estimate (a physical justification must also be document).

    4.2.4   Optionally, compare the variance of the new set of method blanks or FNQS with that of the oiginal (the initial demonstration of performance) using a two-tailed F test

$$F = \frac{\left(s_H^2\right)}{\left(s_L^2\right)}$$

$s_H^2 = \text{Larger variance estimate}$

$s_L^2 = \text{Smaller variance estimate}$

If the calculated value of F is greater than the F-statistic value for the 95% level of confidence (or 99% level may be used when a large number of analytes is being simultaneously determined), then the critical value and detection limit must be updated. If the calculated value of F is less than the F-statistic value for the 95% or 99% level of confidence, then the data should be pooled to calculate a more reliable estimate of L_C. However, it should be noted that the F-test is not robust to departures from normality (e.g., the underlying assumption for the calculation of L_c).[10] .

5.0   DATA REPORTING AND DELIVERABLES

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 11 of 24	Date Initiated:

5.1   List the detection limit L_d on reports. Detections may be reported to the critical value L_c. If detections less than the quantitation limit, L_Q, are required to be reported, then detections greater than or equal to the L_c but less than L_q must be reported as quantitatively estimated values. Report measurements less than L_c as "< L_d" (e.g., or "L_d U," where the "U" qualifier indicates a non-detect). (Note: L_d is the smallest possible reporting limit for non-detects but a larger reporting limit may be used if this meets project objectives.) For methods capable of reporting uncensored results, users may request that results less than L_c be reported (for example, when statistical evaluations will be performed on the data).[13] Reported results less than L_c must be clearly identified as such. One example of appropriate reporting would be to report as "< L_d (Y)", where Y denotes the numerical result obtained from the analytical method.

5.2   Detection limit and critical values should be reported using only one significant figure. For example, if L_d = 1, L_c = 0.6, and the "Limit of Quantitation," L_q, = 2.0; then results are reported as follows (a "J flag" is applied to denote quantitatively estimated results):

Instrument Result	Reported Result
2.1	2.1
1.9	2 J
0.92	0.9 J
0.64	0.6 J
0.38	< 1 or 1U or <1 (0.38)

In this example, the "Limit of Quantitation" (L_Q) is defined as the lowest concentration that, in the context of some tolerance for uncertainty, produces quantitatively reliable results for the end use of the data. (L_q= 2 in this example, for the purposes of illustration only).

5.3   Demonstration of Sensitivity in Various Matrices and for Non-Routine Analytes

5.3.1   For some applications it is not necessary to demonstrate the absolute lowest limit of detection and the minimum requirement of seven replicates may be too onerous. Examples of this situation include:

- When the required reporting level is greater than the laboratory quantitation limit
- When a non-routine analyte is added to a method at the specific request of a client and a higher degree of uncertainty can be tolerated (relative to the degree of acceptable uncertainty for the analytes routinely reported).
- When a client-specific matrix must be evaluated to determine, whether or not the sensitivity of the method in this matrix is adequate or substantially different than the sensitivity in a "clean" matrix such as reagent water.

5.3.2   For these situations analyze a minimum of four method blanks (for uncensored methods) or matrix samples from the same source spiked at or below the level at which sensitivity demonstration is required. (The matrix may be a reference matrix or a matrix applicable to a particular site or project). All blank and spike results must be used for next step.

5.3.3   Calculate the mean and standard deviation of the results. Determine the level of the analyte(s) in method blanks (*e.g.* the mean). Compare the mean and standard deviation with acceptance criteria in the method. If no acceptance criteria are available then recovery must be greater than or equal to 40% and the relative standard deviation must be less than or equal to 50% and the levels in method blanks must be below the required reporting limit or a full detection limit determination according to sections 4.1.1 or 4.1.2 must be performed.

5.3.4   For discerning differences in sensitivity between various matrices the F test may be used as outlined in 4.2.4.[5]

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 12 of 24	Date Initiated:

5.3.5 Gather additional data as the method is performed and calculate the detection limit according to sections 4.1.1 or 4.1.2 once sufficient data is available.

6.0 QUALITY ASSURANCE

Once long-term estimates have been established they should be periodically (at least annually) verified to assure sensitivity has not changed.

6.1 Ongoing verification of L_c using method blanks

6.1.1 Once a large number of points are available (at least 100 data points), at least once per year, evaluate the estimate of L_C using the method blank results. Use all applicable method blanks in the evaluation. Applicable method blanks are all method blanks from analytical batches which were used to report results. Alternatively to an annual review, blanks may be evaluated on an ongoing basis. The number of results greater than L_c should be less than 1%. The probability of obtaining an individual blank result that is greater than L_d is extremely low when the method is in control. A result greater than L_d in the method blank indicates the need for corrective action.

6.1.2 If, for the <u>entire</u> data set of blank results, the proportion of results greater than L_c is more than 2%, then the estimate of L_C is too low and a new estimate must be established.

6.1.2.1 For uncensored methods the most recent 100 method blank results are used to calculate a new L_c.

6.1.2.2 For censored methods that exhibit "sporadic" blank contamination the new L_c is established from the most recent 100 method blank results. As term is is used here, a censored method produces "sporadic" method blank contamination if the analyte is detected in method blanks less than 85% of the time (so that L_C cannot be estimated using the uncensored method approach), but the analyte is detected in the method blanks more than 15% of the time. It is assumed that all efforts have been implemented to minimize blank contamination, prior to operation. Set the new L_c to the value of the largest method blank. However, it should be noted that this estimate of the critical value possesses a somewhat higher level of uncertainty. (This value a non-parametric upper tolerance limit at the 99% level of confidence for approximately 95% rather than 99% coverage; that is, at least 95% of all future method blank measurements should produce analyte concentrations less than this value with 99% confidence.) .

6.2 Onging verification of L_d using the False Negative Quality Control Sample

6.2.1 A false negative quality control check sample (FNQS) must be analyzed periodically for censored and uncensored methods to verify sensitivity and establish the long-term estimates. A FNQS must be analyzed at least monthly or per batch (which ever is less frequent) on each instrument being monitored until such time as 20 applicable data points (including those used to establish the initial estimate) are available and then at least on a quarterly basis thereafter. The FNQS is spiked at or near the detection limit L_d to verify that a measured value greater than L_c is obtained and any method-specified identification criteria are met.

6.2.2 The FNQS is spiked at one to three times the detection limit (L_d) determined from the initial demonstration of performance. The FNQS target concentration is the L_d, however, for multi-analyte methods (i.e. VOCs and SVOCs). A minimum of 90% of the analytes must fall in the 1-3 x L_d spike range and no analyte should be more than five times the L_d. The laboratory must

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 13 of 24	Date Initiated:

routinely report non-detects to values no less than the L_d unless detection capability is demonstrated at lower concentrations (e.g., via the use of lower spike concentrations for the FNQS).

[TG: Still strongly non-concur with this approach: Non-detects should be reported to the FNQS used for on-going demonstration of performance and not the value of L_d determined during initial demonstration of performance when the system is probably at peak performance. A laboratory could fail to detect at L_d while successfully detecting at 3 – 5 L_d due to, for example, a lack of instrument maintenance—This specification is advantageous to analytical testing laboratories, but not data users. It encourages elevated FNQS concentrations and artificially low L_d values---a spike at L_d and not $3L_d$ is needed to verify L_c. The L_d values will already be biased low because most of the L_d determinations will be performed in clean rather than actual environmental sample matrices.

It should also be noted that initial demonstration of performance data are not used in lieu of the on-going performance data for any other QC application. For example, once laboratory control sample (LCS) statistical control limits are determined from the initial demonstration of performance, these limits are updated using LCS recovery data collected as part of the on-going demonstration of performance. The control limits are calculated from the on-going data and not the initial performance data. However, as this procedure is currently written, it allows non-detects to be reported solely upon the basis of the L_d value determined during the initial demonstration of performance; L_d should be strongly dependent upon the on-going performance data like any QC element.]

6.2.3   For on-going verification of the L_d for a specific compound (or analyte) no more than 2% of the FNQS should produce a non-detect (results below the L_c). The 2% failure rate has been exceeded if multiple (two or more) non-detects are produced for the FNQS in a short time period. If the on-going sensitivity verification spike (FNQS) does not produce a detection at least 98% of the time, then perform appropriate instrument maintenance as needed and repeat the FNQS analysis. If detection of the FNQS is reestablished proceed to section 6.2.4. and consider a more frequent routine maintannce schedule. If a detection is still not obtained, method sensitivity may have changed. In all cases, the data set used must be representative of the routine conditions of the preparation and analysis opertions, single injection analysis, with no attempts to optimize (*e.g.* cleaning the MS source, new initial calibration immediately prior, etc.).

6.2.3.1   For uncensored methods calculate a new estimate of L_c and L_d using the most recent 100 blank results. If the new L_d is greater than the FNQS spike concentration this becomes the new L_d. If the new L_d value is less than or equal to the previous L_d analyze a FNQS at a larger concentration (e.g., two times greater than the original FNQS concentration) until a detection (result at or above the L_c) is consistently obtained. Detections must be obtained for at least two consecutive false negative checks at the higher concentration. Use the higher concentration as L_d.

6.2.3.2   For censored methods calculate a new estimate of L_c using the most recent 20-100 FNQS results. If the new L_c value is less than or equal to the previous (section 6.1) L_c this becomes the new L_c. Analyze a FNQS at a larger concentration (e.g., two times greater than the original FNQS concentration) until a detection (at L_c) is consistently obtained. Detections must be obtained for at least two consecutive false negative checks at the higher concentration. Use the higher concentration as L_d.

6.2.4   Whenever major method or instrument changes have occurred, perform the false positive verification test as discussed in section 4.1.2.4. If significant changes in the sensitivity estimate are indicated, repeat the initial detection estimate procedure.

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 14 of 24	Date Initiated:

6.2.5   To most effectively verify L_c and obtain the lowest critical values and detection limits, if significant method bias does not exist, L_d should be no greater than two to three times the calculated value for L_c[11]. Higher values of L_d would be obtained if significant negative bias were to exist (L_d should be no greater than approximately 2 to 3 L_c divided by the mean proportion of recovered analyte). Since meeting this goal for every analyte will not be practical when multiple analytes are simultaneously being quantified (especially when significant bias exists and instrumental response appreciably differs from analytes to analyte or a large list of analyte are being simultaneously quantified), this should be interpreted as *guidance only* (for obtaining the lowest possible critical values and detection limits), rather than as a requirement that must be met for all analytes. For examples of ongoing verification see footnote [12].

_____

7.0   ACKNOWLEDGEMENTS

7.1   The following members of Committee One on Detection contributed to the development of this procedure. Each member represents one or more of the Consensus organizations.

- Steven Bonde – North Creek Analytical Laboratories

- ✓ Richard Burrows, Ph.D. (rburrows@stl-inc.com) – American Council of Independent Laboratories (ACIL), Severn Trent Laboratories (STL)

- Roger Claff, P.E. – American Petroleum Institute (API)

- Nancy Grams – Advanced Earth Technologies (AET), American Society for Testing and Materials D19 (ASTM), Waste Management

- ✓ Thomas Georgian, Ph.D. (Thomas.Georgian@nwd02.usace.army.mil) – United States Army Corps of Engineers (USACE), Standard Methods

- Angie Grooms – Duke Power, Utility Water Act Group (UWAG)

- Donna Hill – Southern Company, Utility Water Act Group (UWAG)

- Larry LaFleur – American Forestry and Paper Association, Inter-Industry Analytical Group (IIAG)

- Patty Lee – Hampton Road Sanitary District (HRSD), Water Environment Federation (WEF)

- ✓ Kenneth Osborn (kosborn@ebmud.com) – East Bay Municipal Utility District (EBMUD), American Water Works Association (AWWA), Standard Methods

- ➢ John Phillips (jphill12@ford.com) – Ford Motor Company, Alliance of Automobile Manufactures (AAM), Inter-Industry Analytical Group (IIAG), Michigan Environmental Laboratories Association (MELA)

- Jim Pletl – Hampton Road Sanitary District (HRSD), Association of Metropolitan Sewerage Agencies (AMSA)

- Lennie Routten – Northrop Grumman, VA American Water Works Association (AWWA), VA Water Environment Association (VWEA)

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 15 of 24	Date Initiated:

- Rock Vitale - Environmental Standards, Inc., representing Environmental Laboratroy Advisory Board (ELAB), American Institute of Chemists (AIC, N National Registry of Cerfified Chemists (NRCC)

➢ *Indicates committee chairman*
✓ *Indicates a major contributor who can be contacted to answer technical questions*

8.0 ATTACHMENTS

8.1 Spreadsheet

8.2 Flowchart
8.3 Grubbs Test

Table 1: K Values

ν	$K_{0.99, 0.01, \nu}$
6	6.101
7	5.529
8	5.127
9	4.829
10	4.599
11	4.415
12	4.264
13	4.138
14	4.031
15	3.939
16	3.859
17	3.789
18	3.726
19	3.67
20	3.619
21	3.573
22	3.532
23	3.494
24	3.458
25	3.426
26	3.396
27	3.368
28	3.342
29	3.317
30	3.295
31	3.273
32	3.253
33	3.234

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 16 of 24	Date Initiated:

34	3.216
35	3.199
36	3.182
37	3.167
38	3.152
39	3.138
40	3.125
41	3.112
42	3.100
43	3.088
44	3.077
45	3.066

ν	$K_{0.99,\,0.01,\,\nu}$
46	3.055
47	3.045
48	3.036
49	3.027
50	3.018
51	3.009
52	3.001
53	2.993
54	2.985
55	2.977
56	2.97
57	2.963
58	2.956
59	2.949
60	2.943
61	2.936
62	2.93
63	2.924
64	2.919
65	2.913
66	2.907
67	2.902
68	2.897
69	2.892
70	2.887
71	2.882
72	2.877
73	2.873
74	2.868
75	2.864
76	2.86
77	2.855
78	2.851

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 17 of 24	Date Initiated:

79	2.847
80	2.843
81	2.839
82	2.836
83	2.832
84	2.828
85	2.825
86	2.821
87	2.818
88	2.815
89	2.811
90	2.808

ν	$K_{0.99,\,0.01,\,\nu}$
91	2.805
92	2.802
93	2.799
94	2.796
95	2.793
96	2.79
97	2.787
98	2.784
99	2.782

For large n (e.g., n> 100),

$$K_{0.99,0.01,\nu} = Z_{0.99}\sqrt{\nu/\chi^2_{0.01}} \approx Z_{0.99}\sqrt{\frac{\nu}{(Z_{0.01}(2\nu)^{1/2} + \nu)}} = 2.326\sqrt{\frac{\nu}{\nu - 2.326(2\nu)^{1/2}}}$$

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 18 of 24	Date Initiated:

Table 2: Student t values

ν	$t_{0.995, \nu}$
6	3.707
7	3.499
8	3.355
9	3.250
10	3.169
11	3.106
12	3.055
13	3.012
14	2.977
15	2.947
16	2.921
17	2.898
18	2.878
19	2.861
20	2.845
21	2.831
22	2.819
23	2.807
24	2.797
25	2.787
26	2.779
27	2.771
28	2.763
29	2.756
30	2.750
31	2.744
32	2.738
33	2.733
34	2.728
35	2.724
36	2.719
37	2.715
38	2.712
39	2.708
40	2.704
50	2.678
60	2.660
70	2.648
80	2.639
90	2.632
100	2.626
150	2.609
200	2.601
Infinite	2.576

[1] If the number of replicates is small, a more reliable estimate of the critical value can be made using the upper tolerance interval (Georgian and Osborn, Quality Assurance, 8:1-9, 2003).

Unless a large number of replicates is available, L_C is calculated a formula of the form:

$$L_c = z_p\, s_{UCL, 1-\gamma} = z_p \sqrt{\nu / \chi^2_{\nu,\,\gamma}}\; s = K_{p,\gamma,\nu}\, s$$

where $\chi^2_{\gamma,\nu}$ is the $\gamma 100^{th}$ percentile of the χ^2 distribution with ν degrees of freedom and z_p denotes the $p100^{th}$ percentile of the standard normal distribution. Note that normality is being assumed. This equation is the $(1-\gamma)$ 100% tolerance interval that contains at least the proportion p of the population. If a large number of replicate measurements were to be performed for a blank, then p100% of the measurements would less than the critical value L_c with $(1-\gamma)$ 100% confidence. If p = 0.99 and γ = 0.01, 99% percent of all future measurements will be less than L_c with 99% confidence. For p = 0.99, γ = 0.01, and ν = 6 (n = 7),

$$L_c = 2.33\sqrt{6/0.872}\; s \; \approx \; 6\,s$$

[2] This is very similar to Currie's critical level, L_C (Anal. Chem. Vol. 40, No. 3 March 1968, p586). It is the level at which a result can be confidently distinguished from the blank. The detection limit, L_D, is set at the lowest value that can be reliably detected, where $L_D \approx 2\,L_c$, assuming constant variance and no analytical bias. Currie's original procedure assumes that the mean result from the blanks will be subtracted from any individual blank result (or equivalently, the absence of significant analytical bias for the final reported result). Since blank subtraction is not allowed in most environmental methods, the mean of the blanks must be added into the calculation.

[3] If spikes at L_D have low average recovery (<70%), then L_D may need to be increased in order to maintain reliable detection (i.e., a result in a measured value greater than L_C). The approximate spiking concentration for L_D can be estimated from the equation:

$$L_D = \overline{X} + 2s\, K_{p,\gamma,\nu}\left(\frac{100}{\%\ \text{Recovery}} \right)$$

[4] If the sample mean ($\overline{X}$) is not significantly different from zero, then the detection limit and critical value may also be determined using the same equation with $\overline{X}$ = 0. It is recommended that a two-tailed t-test be performed at the 99% level of confidence to determine if the mean is significantly different from zero. To do this, calculate the following:

$$t = |\overline{X}| / (s/\sqrt{n})$$

Compare this to the critical value of the Student's t distribution, which is denoted by $t_{1-\alpha/2,\,\nu}$. For a set of n replicate measurements and the $(1-\alpha)$100% level of confidence, this value is the $(1-\alpha/2)100^{th}$ percentile of the Student's t distribution with $\nu = n-1$ degrees of freedom. Table 2 lists the critical values for the Student's t-distribution for the 99% level of confidence (for a two-tailed test) for various degrees of freedom. For n = 7 and α = 0.01, $t_{0.995,\,6}$ = 3.71. If $-t_{1-\alpha/2,\,\nu} < t < t_{1-\alpha/2,\,\nu}$, then there is insufficient evidence to conclude (at the specified level of confidence) that the "true" (population) mean is significantly different from zero and the detection limit and critical value may be calculated by setting the mean equal to zero. The mean is significantly different from zero at the $(1-\alpha)$100% level of confidence, if $|t| > t_{1-\alpha/2,\,\nu}$.

[5] Alternatively, a two-tailed F-test at the 95% or 99% level of confidence may be performed using the highest and the lowest standard deviation to determine if the standard deviations are significantly different from one another.

$$F = \frac{\left(s_H^2\right)}{\left(s_L^2\right)}$$

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 20 of 24	Date Initiated:

$$s_H^2 = \text{Largest variance estimate}$$

$$s_L^2 = \text{Smallest variance estimate}$$

If the calculated value F (above) is less than the critical value of this statistic for the $(1-\alpha)100\% = 95\%$ or 99% level of confidence, $F(n_H-1, n_L-1, 1- \alpha/2)$, then it may be concluded that the instruments possess similar sensitivities and the pooled standard deviation may be used. The values n_H and n_L denote the number of replicates used to calculate the standard deviations s_H and s_L, respectively. If seven replicates are used for all instruments, then the critical values for the 95% and 99% confidence levels are $F (6, 6, 0.975) = 5.82$ and $F(6, 6, 0.995) = 11.07$, respectively. Note that the square root of the critical value is about 2.4 for the 95% level of confidence and 3.3 for the 99% level of confidence. Therefore, if the instruments possess similar sensitivities, the largest standard deviation should not be greater than approximately two or three times the smallest standard deviation. The 95% level of confidence should be used when a small number of analytes are being simultaneously determined (e.g., n < 10); the 99% level of confidence is recommended when a large number of analytes is being determined (e.g., n > 10).

[6] An apparent minimum signal to noise ratio of approximately 3:1 is a reasonable way to evaluate if a measurable signal is achievable for a chromatographic determination. This evaluation need not be done quantitatively, but may be simply estimated visually. For this evaluation integration conditions must be identical to routine operating conditions (e.g. no modification of integration threshold values, peak slope, etc.).

It is critical that the qualitative identification procedures that are used be the same as those employed during routine analysis. Consider the example of a GC/MS test that requires qualifier ions be present in a specific ratio range relative to the primary ion and the identification is rejected if they are not present in that range. In this case it is not acceptable to call the analyte detected for the replicate detection limit study if the qualifier ratios fail. In general, the instrument's target compound identification software should automatically identify the peak without manual intervention to establish reliable qualitative identification. In the case of dual column confirmatory analysis the analyte must be detected on both columns. In the case of pattern recognition, for example Aroclor analysis, the analyst must be able to recognize the pattern.

[7] It is recognized that x will probably not be the same for each analyte in a multi-analyte test.

[8] For some isotope dilution methods an Estimated Detection Limit (EDL) may be measured and calculated for each analytical result. The EDL reflects the sample concentration of the particular analyte which would be required to cause a peak at 2.5 times the background noise (i.e. a positive result) for the particular analysis. The EDL is calculated according to the following equation:

$$\text{Estimated Detection Limit} = \frac{N \times 2.5 \times Qis}{His \times RRF \times W \times S}$$

where:

N	=	peak to peak noise of quantitation ion signal in the region of the ion chromatogram where the compound of interest is expected to elute
His	=	peak height of quantitation ion for appropriate internal standard
Qis	=	ng of internal standard added to sample
RRF	=	mean relative response factor of compound obtained during initial calibration
W	=	amount of sample extracted (grams or liters)
S	=	percent solids (optional, if results are requested to be reported on dry weight basis)

If this procedure is used, determination of the detection limit is the higher of the EDL or L_d.

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 21 of 24	Date Initiated:

[9] Where the population sample has more than one method blank value exceeding L_c, L_c is suspect. Where more than two method blanks of the last 100 exceed L_c the laboratory may either increase L_c by one unit (e.g., 0.04 to 0.05) or it may repeat the estimation procedure. Continue to monitor and adjust L_c as needed.

[10] Due to departures from normality, it could be concluded that the variances are significantly different when they are not. Therefore, it is strongly recommended that Levene's test be used in lieu of the F test to determine if the standard deviations (variances) differ at 95% or 99% level of confidence.

[11] A two-tailed t-test at the 99% level of confidence may be performed to determine if the mean percent recovery is significantly different from 100%. To do this, calculate the following:

$$t = (100 - \overline{X}) / (s / \sqrt{n})$$

Compare this to the critical value of the Student's t distribution, which is denoted by $t_{1-\alpha/2, \nu}$. For a set of n replicate measurements and the $(1-\alpha)100\%$ level of confidence, this value is the $(1-\alpha/2)100^{th}$ percentile of the Student's t distribution with $\nu = n - 1$ degrees of freedom. Table 2 lists the critical values for the Student's t-distribution for the 99% level of confidence (for a two-tailed test) for various degrees of freedom. For n = 7 and $\alpha = 0.01$, $t_{0.995, 6} = 3.71$. If $-t_{1-\alpha/2, \nu} < t < t_{1-\alpha/2, \nu}$, then there is insufficient evidence to conclude (at the specified level of confidence) that the mean recovery is significantly different from 100% (i.e., there is insufficient evidence to conclude bias exists). The mean recovery is significantly different from 100% (i.e., bias exists) at the $(1-\alpha)100\%$ level of confidence, if $|t| > t_{1-\alpha/2, \nu}$.

[12] Examples

Example 1

X = mean result,   SD = Standard deviation

Detection limits determined on the basis of method blanks:

<u>Original short term</u> (7 replicate)

Standard deviation = 2.5   K = 6.1   Mean = 0

$Lc = X + K*SD_{Blanks} = 6.1 * 2.5 = 15$

$Ld = X + 2K* SD_{Blanks} = 30$

<u>Long term data</u>

FNQS concentration = 40, Range of FNQS results = 25-55

$SD_{FNQS} = 5$ Number of FNQS = 20

$SD_{Blanks} = 4$ $X_{Blanks} = 1$ Number of blanks = 100

$Lc = X + K*SD_{Blanks} = 1 + 2.8*4 = 12$

Ld is not adjusted since results are above Lc. It may be possible to document that Ld is lower than 40. If the lab wants to do this then data with a spike level lower than 40 would have to be collected.

<u>Ongoing verification</u>

The lab verifies that less than 2% of blank results are above 12.

Example 2

Detection limits determined on the basis of method blanks:

<u>Original short term</u> (7 replicate)

Standard deviation = 2.5,   K = 6.1   Mean = 0

$Lc = X + K*SD_{Blanks} = 6.1 * 2.5 = 15$

$Ld = X + 2K* SD_{Blanks} = 30$

<u>Long term data</u>

FNQS concentration = 40   Range of FNQS results = 15-55

$SD_{FNQS} = 10$ Number of FNQS = 20

$SD_{Blanks} = 8$ $X_{Blanks} = 1$ Number of blanks = 100

$Lc = X + K*SD_{Blanks} = 1 + 2.8*8 = 23$

Ld is adjusted since some FNQS results are below Lc. Based on the standard deviation of the FNQS results Ld would need to be raised to at least 50 for 99% of results to be above the Lc value of 23

<u>Ongoing verification</u>

The lab verifies that less than 2% of blank results are above 23

The lab verifies that less than 2% of FNQS results at the new spike level (50) are below 23

Example 3

Detection limits determined on the basis of spikes

Original short term (7 replicate)

Spike concentration 30

Standard deviation of spikes = 2.5   K = 6.1

$Lc = +K*SD_{Spike} = 6.1 * 2.5 = 15$

No spike results are < 15

$Ld = 30$

Long term data

FNQS concentration = 30   Range of FNQS results = 15-55

$SD_{FNQS} = 6$, Number of FNQS = 20

$Lc = K*SD_{Spikes} = 3.67*6 = 22$

Lc is adjusted to 22

Some FNQS results are below 22. Therefore FNQS and Ld are raised to 44 (Concentration expected to return results above 22 at least 99% of the time, based on the standard deviation)

Ongoing verification

The lab verifies that less than 2% of FNQS results at the new spike level (44) are below 22

Example 4

Detection limits determined on the basis of spikes

Original short term (7 replicate)

Spike concentration 30

Standard deviation of spikes = 2.5   K = 6.1 Mean recovery =50% (mean result is 15)

$Lc = K*SD_{Spike} = 6.1 * 2.5 = 15$

Many spike results are < 15

Spike concentration is raised to 45 Ld is set at 45 (K * SD * 1/Recovery + Lc) =[(6.1 * 2.5 * 2) + 22] = 66

Long term data

FNQS concentration = 45   Range of FNQS results = 4.5 - 36

$SD_{FNQS} = 6$ Number of FNQS = 20

$Lc = K*SD_{FNQS} = 3.67*6 = 22$

Lc is adjusted to 19

Some FNQS results are below 22. Therefore FNQS and Ld are raised to 66 (Concentration expected to return results above 19 at least 99% of the time, based on the standard deviation and 50% recovery of the spikes. The calculation for the desired spiking level is (K * SD * 1/Recovery + Lc = [(3.67 * 6 * 2) + 22] = 66

Ongoing verification

The lab verifies that less than 2% of FNQS results at the new spike level (66) are below 19

SOP Name:	**Proposed Procedures for Estimating the Limit of Detection**	Revision Number: 0
		Date Revised:
SOP Number:	page 24 of 24	Date Initiated:

[13] The International Union of Pure and Applied Chemistry (IUPAC) recommendation is to "…*always* report both the estimated value of the measured quantity ($\hat{L}$) and its uncertainty, *even* when $\hat{L} < L_C$ results in the decision 'not detected.' Otherwise, there is needless information loss, and, of course, the impossibility of averaging a series of results." (*Chemometrics and intelligent laboratory systems* (Currie, 1997), p. 156). Reporting below Lc may not be practical, however, and reported values below Ld would be inappropriate for making compliance determinations. However, measurement values below Lc can be used in statistical evaluations of large data sets for the purpose of calculating means, standard deviations, confidence intervals, and future event likelihoods. Examples of such use would include the comparison of two analytical methods in the region of analytical lower limit capability, determination of a background contribution allowance, and reasonable potential analysis.